Annotate and Identify Modalities, Speech Acts and Finer-Grained Event Types in Chinese Text

Hongzhi Xu
Department of CBS
The Hong Kong Polytechnic University
hongz.xu@gmail.com

Chu-Ren Huang
Faculty of Humanities
The Hong Kong Polytechnic University
churenhuang@gmail.com

Abstract

Discriminating sentences that denote modalities and speech acts from the ones that describe or report events is a fundamental task for accurate event processing. However, little attention has been paid on this issue. No Chinese corpus is available by now with all different types of sentences annotated with their main functionalities in terms of modality, speech act or event. This paper describes a Chinese corpus with all the information annotated. Based on the five event types that are usually adopted in previous studies of event classification, namely state, activity, achievement, accomplishment and semelfactive, we further provide finer-grained categories, considering that each of the finer-grained event types has different semantic entailments. To differentiate them is useful for deep semantic processing and will thus benefit NLP applications such as question answering and machine translation, etc. We also provide experiments to show that the different types of sentences are differentiable with a promising performance.

1 Introduction

Event classification is a fundamental task for NLP applications, such as question answering and machine translation, which need deep understanding of the text. Previous work (Siegel, 1999; Siegel and McKeown, 2000; Palmer et al., 2007; Zarcone and Lenci, 2008; Cao et al., 2006; Zhu et al., 2000) aims to classify events into four categories, namely state, activity, accomplishment and achievement, i.e. Vendler's framework adopted from linguistic studies (Vendler, 1967; Smith, 1991). High performance was reported on the classification, however based on the assumption that all sentences describe an event, which is not case in real text. Modalities and speech acts are not considered and no finer-grained classification is proposed.

The aim for aspectual classification for a specific language is to build verb classes. In such framework, viewpoint aspect in terms of perfective vs. imperfective is not considered. For example, *he is eating a sandwich* and *he ate a sandwich* are all instances of accomplishment. However, we argue that this framework is not enough for more accurate event processing. It is obvious that the two sentences have different meanings and different consequences. The situation described by the first sentence is still going on at the speech time, while the second sentence implies that the event has finished. So, in the perspective of event processing, it is necessary and important to discriminate the two different aspects.

Another important issue is that not all sentences describe events. For example, Austin (1975) discriminated two different types of sentences: constative and performative. Sentences that report or describe events are in the first category. Sentences of the performative category mainly refer to speech (illocutionary) acts, actions that are done by speech. For example, by uttering the sentence *I declare that the new policy will take effect from now on*, the authorized speaker brings a new policy into effect. In this case, uttering the sentence itself is an event. Discriminating speech acts are especially useful in speech corpora, e.g. (Avila and Mello, 2013).

Modality is important due to its interaction with factuality and truth of the embedded propositions. For example, *he can eat two sandwiches* describes a dynamic modality about the subject's ability of eating.

Proceedings of the Workshop on Lexical and Grammatical Resources for Language Processing, pages 157–166,
Coling 2014, Dublin, Ireland, August 24 2014.

However, no eating event has actually happened. Modality has been considered in modeling speaker's opinions (Benamara et al., 2012), machine translation (Baker et al., 2012), etc.

Sauri et al. (2006; 2012) proposed a framework for modeling modalities. However, their definition of modality is a little different from that used by linguists. The main motivation of their work is to predict the factuality of a proposition. As a result, all factors that may affect the factuality of propositions are regarded as modalities. In our framework, we will adopt the definition in linguistic studies that modality expresses a speaker's belief or attitude on an embedded proposition (Palmer, 2001). Factuality is determined by many factors other than modalities. However, we don't want to mix all the factors together in linguistic perspective.

In this paper, we will describe a Chinese corpus in which different sentence types are discriminated. Finer-grained event types are also incorporated with a theory proposed in (Xu and Huang, 2013). The details of the framework will be discussed in the next section.

The remaining of the paper is organized as follows. Section 2 introduces the theoretical framework we shall adopt for our annotation. Section 3 describes a Chinese corpus we annotated with some statistical information. Section 4 describes a classification experiment based on the annotated corpus. Section 5 is the conclusion and our future work.

2 The Annotation Framework

In this section, we will give an introduction to the theoretical framework from a linguistic perspective. There are two main levels for the classification. Sentences are first discriminated according to their main functions, e.g. constative and performative (Austin, 1975). Constative sentences are further divided into modality which mainly expresses the addresser's propositional attitude and event which is a description or report of a real situation without the speaker's attitude. One basic assumption is that one sentence only has one main function in terms of expressing speaker's modality, speech act or describing an event. So, there is no overlap among the three types of sentences.

2.1 Modality

Sentences denoting modalities are different from the sentences reporting events in that the former only refers to a proposition upon which the speaker expresses his attitude its truth value, while the later is a fact without incorporating speakers' opinions but only speaker's perception. It is possible that speakers can make mistakes in their perceptions. However it is beyond the linguistic level and there is no way to predict the correctness based on the surface of the sentence. Thus, it is another issue out of the discussion of this paper. We adopt the modal theory by Palmer (2001). According to him, modality could be divided into epistemic, deontic and dynamic.

Epistemic modality expressed the speaker's opinion on the truth of the embedded proposition in terms of necessity and possibility. Informally, epistemic modality expresses what may be in the world. For example, *ta1 ken3ding4 zai4 ban4gong1shi4* "he must be in his office" describes an epistemic modality of the speaker that he is sure about the truth of the embedded proposition.

Deontic modality expresses what should be in the world, according to speaker's expectations, certain rules, laws and so on. For example, *ni3 bi4xu1 zun1shou3 gui1ze2* "You must obey the rules".

Dynamic modality describes the abilities of a subject, such as *ta1 hui4 you2you3* "he can swim", *wo3 de0 ban4gong1shi4 ke3yi3 kan4jian4 da4hai3* "you can see the ocean from my office".

Evaluation is also treated as a modality in our framework. Evaluation describes the speaker's opinion on a proposition. It is different from epistemic in that it suggests rather than makes judgment on the truth of a proposition. For example, *ta1 suan4shi4 shi4jie4shang4 zui4hao3 de0 ge1shou3 le0* "he should be the best singer in the world". Evaluative sentences only refer to those that contain explicit markers, e.g. *suan4shi4* "should be". The sentence *ta1 shi4 shi4jie4shang4 zui4hao3 de0 ge1shou3* "he is the best singer in the world" is not treated as evaluation. In this sense, evaluative is not equivalent to subjective.

Exclamation is treated as a subset of evaluation. Take *nian2qing1 ren2 a0 !* “Young people!” for example, it mostly expresses an implicit evaluation, e.g. only young people could do crazy things of some kind, based on which the exclamation is expressed by the speaker.

2.2 Speech act

For speech act (illocutionary act), we adopt the theory by Searle (1976), where five different categories are proposed, namely assertive, expressive, directive, commissive and declaration. In addition, we also put interrogative sentences under this category. Speech act sentences only refer to those sentences that are explicit utterances, e.g. the sentences quoted in text.

Assertive is to commit the speaker (in varying degrees) to something’s being the case or the truth of the expressed proposition. For example, *wo3 zheng4ming2 ta1 shi4 xue2sheng1* “I certify that he is a student”.

Expressive expresses the psychological state specified in the sincerity condition about a state of affairs specified in the propositional content. Verbs for expressive speech act includes *xie4xie4* “thank”, *bao4qian4* “apologize”, *huan1ying2* “welcome”, *dui4bu4qi3* “sorry” etc. For example, *xie4xie4 bang1mang2* “Thanks for your help”.

Directive is usually a command or requirement of the speaker to get the hearer to do something. For example, *ni3 guo4lai2 yi1xia4* “Come here please”.

Commissive is to commit the speaker (in varying degrees) to some future course of action. For example, *wo3 hui4 bang1 ni3* “I shall help you”.

Declaration is to bring about the correspondence between the propositional content and reality. Successful performance guarantees that the propositional content corresponds to the world. For example, *wo3 xuan1bu4 ben3 ci4 hui4yi4 zheng4shi4 kai1mu4* “The conference now start”.

Interrogative is an illocutionary act of the speaker that requires the hearer to provided some information. For example, *ni3 jiao4 shen2me0 ming2zi4 ?* “What’s your name?” and *ni3 qu4 ting1 na4 ge4 jiang3zuo4 ma0 ?* “Will you attend the speech?” Interrogative sentences are usually with a question mark “?”. However, not all sentences with question mark are interrogative. For example, rhetorical questions usually don’t need the answer from the hearer. Instead, it actually expresses the speaker’s evaluation on a situation. For example, the sentence *wo3 zen3me0 ke3yi3 bu4 jin4xin1 zhao4gu4 ?* “How could I not take care of him carefully?” should be labelled as evaluative modality rather than interrogative speech act.

2.3 Events

Here, we describe a new framework by incorporating finer-grained event categories as described in (Xu and Huang, 2013). Each of the finer-grained categories corresponds to only one of the five coarse categories. So, it is an extension of and is compatible with the Vendler’s framework.

2.3.1 Primitive Events

According to Xu and Huang (2013), there are three event primitives, namely static state (S), dynamic state (D), and change of state. Static state is equivalent to the previous notion *state*, which is a homogeneous process, where all subparts are of the same kind of event. Dynamic state refers to an ongoing dynamic process, e.g. running, eating etc., that is perceived like a state. Change of state is then defined as a change from one state, either static or dynamic, to another state.

Change of state actually refers to the previous notion *achievement*. Theoretically, there are four types of changes: static-static change (SS), static-dynamic change (SD), dynamic-static change (DS) and dynamic-dynamic change (DD). In detail, SD change is somewhat equivalent to inceptive achievement, and DS change is somewhat equivalent to terminative or completive achievement.

Event Type	Representation	Example	
Static State	—–	ta1 hen3 gao1	*he is tall*
Dynamic State	˜˜˜˜	ta1 zai4 pao3bu4	*he is running*
SS Change	— | —	ta1 bing4 le0	*he got ill*
SD Change	— | ˜˜˜	ta1 kai1shi3 pao3bu4 le0	*he started running*
DS Change	˜˜˜ | —	ta1 ting2zhi3 pao3bu4 le0	*he stopped running*
DD Change	˜˜˜ | ˜˜˜	dian4nao3 qi3dong4 hao3 le0	*the computer finished startup*

Table 1: Primitives of Events.

Table 1 shows the extended event primitives with some illustrative examples. We use '—' and '˜˜˜' to denote static state and dynamic state respectively. ' | ' is used to denote a temporal boundary. In case of change of state, the temporal boundary overlap with the logical boundary, i.e. the change.

Negations usually denote static state. In Chinese, there are two negation adverbs, *bu4* "not" and *mei2you3* "not". However, they are different in that the former negates a generic event meaning that such event doesn't happen, while the latter negates the existence of an event instance. For example, *ta1 bu4 he1jiu3* "he doesn't drink" describes an attribute of the subject, which is intrinsically a static state. *ta1 mei2you3 he1jiu3* "he didn't drink" describes a fact that there is no event instance of his drinking, which is also a static state. Negation of a modality is still a modality. For example, *ta1 bu4 ke3neng2 zai4 ban4gong1shi4* "he cannot be in his office" still describes an epistemic modality.

2.3.2 Complex Events

Based on the primitives, we can compose complex events. Delimitative describes a temporal bounded static state that has a potential starting point and ending point, within which the static state holds, e.g. *ta1 bing4 le0 yi1 ge4 xing1qi1* "he was ill for one week". Process describes a temporal bounded dynamic state that has a potential starting point and ending point, within which the dynamic state holds, e.g. *ta1 pao3 le0 yi1 ge4 xiao3shi2* "he ran for one hour". Semelfactive is different from Process in that its durations is quite short and is usually perceived as instantaneous. In other words, the temporal boundaries of semelfactive is usually naturally determined. For example, *ta1 qiao1 le0 yi1 xia4 men2* "he knocked the door once". There is no way to length the duration of the knocking action. However, a series of iterative semelfactives could form dynamic process. For example, *ta1 qiao1 le0 yi1 ge4 xiao3shi2 de0 men2* "he knocked the door for an hour" gives a reading of iterative knocks.

For static state and dynamic state, we can only refer to their holding at a certain time point. In other words, delimitative and process describe the life cycle of a state. For example, *ta1 bing4 zhe0 ne0* "he is ill" and *ta1 wan3shang4 jiu3dian3 de0 shi2hou0 zai4 pao3bu4* "He was running at 9:00pm". It is also possible to claim that in a certain period, which for some reason became the focus of a conversation, a state holds. For example, *ta1 na4 liang3 tian1 dou1 bing4 zhe0* "he was ill in that two days" and *ta1 wan3shang4 jiu3dian3 dao4 shi2dian3 de0 shi2hou0 zai4 pao3bu4* "From 9:00pm to 10:00pm, he was running". In this case, they are also state rather than delimitative or process. The difference is that there is no information about the starts and the ends, while delimitative and process do.

Accomplishment is composed by a process with a final state. For example, *ta1 xie3 le0 yi1 feng1 xin4* "he wrote a letter" describes an accomplishment composed by a writing process with a final state, i.e. the existence of the letter. The final state of an accomplishment could also be dynamic. For example, *ta1 ba3 dian4nao3 qi3dong4 le0* "he started up the computer" describe an accomplishment with a dynamic final state, i.e. the normal working of computer.

Some Resultative Verb Compounds (RVCs) in Chinese can denote achievements. However, they are easy to be confused with accomplishment. Based on the representation, the difference of them is that accomplishment encodes the start of the dynamic process, while achievement doesn't. For example, *ta1 xie3 wan2 le0 na4 feng1 xin4* "He (write-)finished the letter" describes a DS change. To differentiate them, we can use the *yi3qian2* "before" test. As in this example, *ta1 xie3 wan2 na4 feng1 xin4 yi3 qian2* "before he finished the letter" refers to the period that includes the writing process. This means that

RVCs only focus on the final culminating point and are thus achievements. On the other hand, *ta1 xie3 na4 feng1 xin4 zhi1 qian2* "before he wrote the letter" refers to the period before the writing process. So, *ta1 xie3 le0 yi1 feng1 xin4* "he wrote a letter" is then an accomplishment.

There is a counterpart for accomplishment, which is composed by an instantaneous dynamic process (semelfactive) with a final state. RVCs can also denote instantaneous accomplishment. For example, *ta1 da3sui4 le0 yi1 ge4 bei1zi0* "he hit and broke a cup" is an accomplishment composed by a semelfactive hitting action with a final state, i.e. the broken of the cup. Similarly, the final state could also be dynamic. For example, in *ta1 tan2zhuan4 le0 yi1 ge4 shai3zi0* "He flicked and putted a spin on the dice", the predicate *tan2zhuan4* "flick-spin" is a compound that combines the predicate *tan2* "flick" and *zhuan4* "spin". The whole event is composed by a semelfactive flicking and a final dynamic state of the dice's spin.

Table 2 shows the seven event types with examples. Theoretically, there could be unlimited number of complex events. However, the notions listed here are important in that they are the lexicalized units which reflect the human's cognition of real world events. For the perspective of computational linguistics, discriminating all these linguistic events will be a fundamental step for deeper natural language understanding.

2.3.3 The Neutral Aspect

Some sentences don't include an explicit viewpoint aspect, e.g. without any aspectual markers. For example, *ta1 kan4 xiao3shuo1* "he read novel" can possibly denote different event types in different contexts. *yi3qian2, ta1 kan4 xiao3shuo1* "he read novel before" denotes an attribute of the subject that he reads novels, while *da4jia1 dou1 hen3mang2, xiao3hai2er0 xie3 zuo4ye4, ta1 kan4 xiao3shuo1* "Everyone is busy, children are doing homework, he is reading novels" describes a dynamic state. The aspects of these examples are given by the specified contexts. Such sentences are usually called with NEUTRAL aspect (Smith, 1991). In our framework, such sentences are ignored for now, unless the context can help the annotator to figure out the aspectual information.

<table>
<tr><td>Semelfactive</td><td>|˜|</td><td>ta1 qiao1 le0 qiao1 men2</td><td>"he knocked the door"</td></tr>
<tr><td>Delimitative</td><td>|—-|</td><td>ta1 bing4 le0 yi1 ge4 xing1qing1</td><td>"he was ill for one week"</td></tr>
<tr><td>Process</td><td>|˜˜˜˜|</td><td>ta1 pao3 le0 yi1 ge4 xiao3shi2</td><td>"he ran for an hour"</td></tr>
<tr><td rowspan="2">Instantaneous
Accomplishment</td><td>|˜|—</td><td>ta1 da3sui4 le0 bei1zi0</td><td>"he broke the cup"</td></tr>
<tr><td>|˜|˜˜˜</td><td>ta1 tan2zhuan4 le0 yi1 ge4 shai3zi0</td><td>"He putted a spin on the dice"</td></tr>
<tr><td rowspan="2">Accomplishment</td><td>|˜˜˜˜|—</td><td>ta1 xie3 le0 yi1 feng1 xin4</td><td>"he wrote a letter"</td></tr>
<tr><td>|˜˜˜˜|˜˜˜</td><td>ta1 ba3 dian4nao3 qi3dong4 le0</td><td>"he started up the computer"</td></tr>
</table>

Table 2: Complex event types that are composed by more than one primitives.

The overall hierarchy is shown in Figure 1. Some traditional notions are kept in use e.g. *accomplishment* and *achievement*. However, they now refer to event types rather than verb classes.

3 Annotating a Chinese Corpus

3.1 Data Selection

For annotation, we choose Sinica Treebank 3.0 (Huang et al., 2000), which contains more than 60,000 trees. Sinica Treebank is a subset of Sinica Corpus (Chen et al., 1996), which is a balanced corpus that contains different genres of materials, including news, novels and some transcripts of spoken Chinese. Sinica Treebank is annotated based on the Information-based Case Grammar (Chen and Huang, 1990). The annotated syntactic and semantic information is kept for further studies, e.g. feature evaluation and selection.

For annotation, we only select the sentences that are labeled as S and end with punctuation of period '。', exclamation '！', semicolon '；' and question mark '？'. After removing duplicate sentences, we get 5612 sentences Table 3 shows the detailed information of the raw corpus. There are 45728 tokens from 11681 types in the corpus. For the heads of the sentences, there are 2127 different verbs.

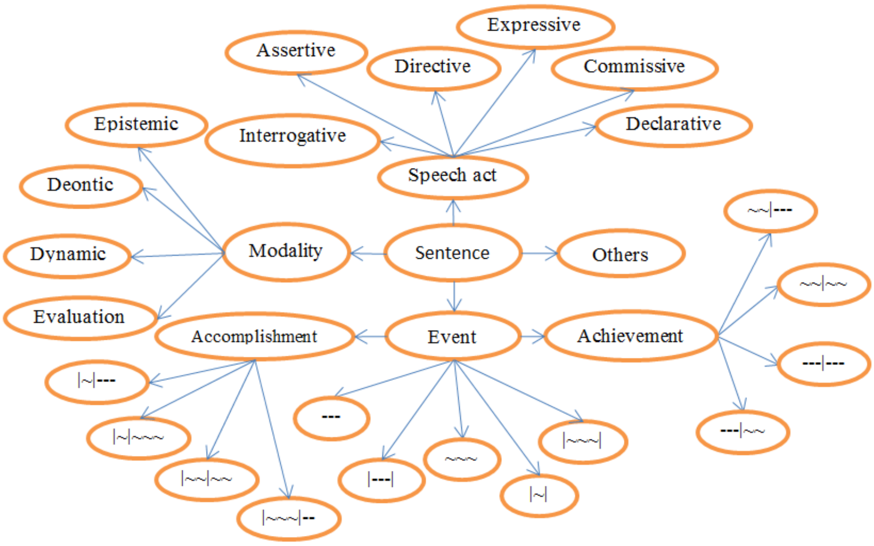

Figure 1: Sentence type hierarchy.

Sentences	Different Verbs	Different Words	Tokens	Characters
5612	2127	11681	45728	75960

Table 3: Distribution information of the corpus for annotation.

3.2 Annotation Result

Each sentence is labeled as one specific finer-grained category from the 23 categories described in Section 2. Whenever an example could not be decided by the annotator, it is discussed with another two linguistic experts to make the final decision. However, we also did agreement test, which will be discussed later.

Finally, we annotated 1044 instances in modality, 764 speech act instances and 3811 event instances. The distribution information is shown in Table 4. We can see that some event types, although theoretically exist, don’ t encounter any examples, such as the instantaneous accomplishment with dynamic final state: |˜|˜˜˜.

Static state contains more than 40% instances. We think that it reflects the real distribution of event types as we don’ t make any bias for selecting data. Static state can be further divided into several subcategories, e.g. attributive, relational, habitual, etc., which will be our future work.

<table>
<tr><th>Type</th><th>No.</th><th>Type</th><th>No.</th><th>Type</th><th>No.</th><th>Type</th><th>No.</th><th>Type</th><th>No.</th></tr>
<tr><td>Epistemic</td><td>303</td><td>Assertive</td><td>64</td><td>—</td><td>2475</td><td>—|—</td><td>471</td><td>|˜|—</td><td>257</td></tr>
<tr><td>Deontic</td><td>219</td><td>Expressive</td><td>13</td><td>˜˜˜</td><td>166</td><td></td><td></td><td>|˜|˜˜˜</td><td>0</td></tr>
<tr><td>Dynamic</td><td>111</td><td>Directive</td><td>65</td><td>|—|</td><td>6</td><td>—|˜˜˜</td><td>96</td><td>|˜˜˜˜|—</td><td>163</td></tr>
<tr><td>Evaluation</td><td>411</td><td>Commissive</td><td>58</td><td>|˜˜˜|</td><td>48</td><td>˜˜˜|—</td><td>79</td><td>|˜˜˜˜|˜˜˜</td><td>40</td></tr>
<tr><td>Interrogative</td><td>559</td><td>Declarative</td><td>2</td><td>|˜|</td><td>4</td><td>˜˜˜|˜˜˜</td><td>2</td><td></td><td></td></tr>
</table>

Table 4: Distribution of different event types in the annotated corpus.

Table 5 shows the number of the main verbs regarding how many event types they can denote excluding modality and speech act. We can see that more than 200 verbs correspond to more than one category. This shows that the verbs alone sometimes could not determine the event type.

No. of Event Types	1	2	3	4	5	6	7
No. of Verbs	1395	155	44	9	7	1	1

Table 5: Number of verbs with regard to how many event types they can denote.

	Accuracy	F1-Measure	Kappa
Annotator 1	0.862	0.762	0.837
Annotator 2	0.821	0.677	0.784
Annotator 1+2	0.842	0.716	0.811

Table 6: Annotation agreements between the main annotator and annotator1, annotator 2, annotator 1+2. Annotator 1+2 means the combination result of the two annotators, i.e. all the 2000 examples.

3.3 Agreement Evaluation

In order to test the reliability of the annotation, we randomly select 2000 examples from the corpus and let another two linguists annotate them. Each of the linguists annotate half of them. The annotation results are then compared with the main annotator. The agreements between the main annotator and the other two annotators in terms of accuracy, F1 measure and Kappa value are shown in Table 6. The F1 measures are calculated based on the assumption that the main annotator's result is the gold standard. The result shows a very high agreement which means that our new framework for event type classification is reliable and easy for annotation.

4 Automatic Classification of Chinese Sentences and Event Types

In this section, we conduct two classification experiments. The first is to discriminate the three sentence types regarding their main functions, speech act, modality and event. The second is the classification with the finer-grained categories. Before the experiments, we will first discuss the features that may help for the classification.

4.1 Features

As suggested in previous literatures (Siegel, 1999; Siegel and McKeown, 2000; Zhu et al., 2000; Cao et al., 2006), the following features are considered as important for event type classification.

Main verbs and their complements including argument structure are the most important indicators to an event type. Negation of the main verb is a strong indicator for static state, as discussed above.

Aspectual markers, 着 *zhe0* "ZHE", 了 *le0* "LE", 过 *guo4* "GUO" and some aspectual light verbs, e.g. 在 *zai4* "be doing", 开始 *kai1shi3* "start", 继续 *ji4xu4* "continue", 停止 *ting2zhi3* "stop", 完成 *wan2cheng2* "finish", are strong indicators for different event types.

Temporal adverbials are also important features, which could potentially disambiguate neutral sentences, e.g., *yi3qian2, ta1 kan4 xiao3shuo1* "he read novel before" as discussed above.

Frequency adverbs, such as 经常 *jing1chang2* "often", 偶尔 *ou3er3* "sometimes", etc., are indicators for habitual states. For example, *ta1 jing1chang2 qu4 he1jiu3* "he often goes for drinking" is a habitual state rather than a specific event.

Modalities could be expressed by auxiliaries, adverbs, sentence final particles etc. in Chinese. Adverbs that modify the main verb, such as 可能 *ke3neng2* "possibly", are important features for identifying modalities. Sentence final particles (SFP) and punctuation marks are also good indicators to evaluative modality.

Since we don't maintain a dictionary for the above indicators, we use a general feature set including the dependency structure and the combinations of the dependent constituents. We suggest that the above linguistic rules could be reflected by the dependency structures, which could be captured by the classifiers. Meanwhile, the experiment result here is only to serve as a baseline for future comparisons. In all, the features are listed in Table 7 with some examples.

ID	**Feature**	**Example**
f_1	Head	head:word:kan4, head:pos:verb, head:subj:word:ta1, head:subj:pos:pron, head:obj:xp:NP, head:obj:xp:noun-noun
f_2	Dependency	dep:word:ta1, dep:pos:pron, dep:word:bu4, dep:pos:adv, dep:word:xiao3shuo1, dep:pos:noun, dep:word:le0, dep:pos:particle,
f_3	COMB	subj:word:ta1-head:word:kan4-obj:xp:noun-noun, subj:pos:pron-head:pos:verb-obj:xp:NP,

Table 7: Feature template we use for our classification of event types. Feature examples are based on the sentence *ta1 (he) bu4 (not) kan4 (read) zhen1tan4 (detective) xiao3shuo1 (novel) le0 (LE)* "he doesn't read detective novels any more".

	f_1			$+f_2$			$+f_3$		
	Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1
Event	0.709	0.939	0.807	0.853	0.969	0.908	0.833	0.974	0.898
Modality	0.395	0.124	0.189	0.731	0.473	0.574	0.744	0.431	0.545
SpeechAct	0.430	0.130	0.199	0.829	0.664	0.737	0.845	0.609	0.707
MacroAvg	0.511	0.398	0.399	0.804	0.702	0.740	0.807	0.671	0.717
Accuracy	0.679			0.836			0.824		

Table 8: Coarse level classification result.

4.2 Experimental Result

To give a real performance, the annotated syntactic and semantic information are not used. Instead, we use the Stanford word segmenter (Tseng et al., 2005) and Stanford parser (Chang et al., 2009) to get the syntactic structure of the sentences. All the experiment are results of 5-fold cross validation with a SVM classifier implemented in LibSVM (Chang and Lin, 2011).

The result of the coarse level classification for modality, event and speech act is shown in Table 8. We can see that the overall performance is reasonable. The F-Measure for modality is not as good as the others. This is due to the fact that the modal markers and operators are quite critical for identifying modalities, which may be sparse in our corpus. We suggest that maintaining a comprehensive dictionary of modal operators could benefit the identification of the modalities. We can also see that the feature set f_3 harms the performance, which is also caused by the feature sparseness problem.

For finer-grained classification, we use two different ways. The first way is to use a hierarchical classification scheme. An instance is first classified as event, modality or speech act. According to the result of the first round classification, the instance is put into the corresponding finer-grained model for further classification. The second way is to classify all instances all at once based on a model trained on all finer-grained categories.

Considering that some categories contain only few examples, which will provide unreliable evaluation of the performance, we combined accomplishments with static final state and dynamic state, so does for instantaneous accomplishment. We use '=' to denote a general state, which could be either static or dynamic. Static state and delimitative are combined together, while dynamic state, process and semelfactive are combined. Expressive, declarative and DD change are ignored in the experiments. The classification results with feature sets f_1 and f_2 are shown in Table 9. The hierarchical classification is slightly better than the all-at-once classification. Meanwhile, the accuracy for hierarchical classification is 0.621, which is much better than the predominant guess 0.443.

We should note that parsing accuracy will significantly affect the result of event type classification. This is true in the sense that the semantic content of words and their syntactic relations are all critical

<table>
<tr><th></th><th colspan="3">All-At-Once</th><th colspan="3">Hierarchical</th></tr>
<tr><th></th><th>Precision</th><th>Recall</th><th>F1</th><th>Precision</th><th>Recall</th><th>F1</th></tr>
<tr><td>—</td><td>0.609</td><td>0.952</td><td>0.743</td><td>0.627</td><td>0.938</td><td>0.751</td></tr>
<tr><td>~~~</td><td>0.840</td><td>0.078</td><td>0.142</td><td>0.830</td><td>0.069</td><td>0.127</td></tr>
<tr><td>—|—</td><td>0.454</td><td>0.384</td><td>0.415</td><td>0.473</td><td>0.418</td><td>0.443</td></tr>
<tr><td>—|~~~</td><td>0.583</td><td>0.083</td><td>0.142</td><td>0.537</td><td>0.104</td><td>0.173</td></tr>
<tr><td>~~~|—</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>|~~~~|===</td><td>0.438</td><td>0.084</td><td>0.140</td><td>0.394</td><td>0.108</td><td>0.168</td></tr>
<tr><td>|~|===</td><td>0.496</td><td>0.159</td><td>0.239</td><td>0.516</td><td>0.210</td><td>0.295</td></tr>
<tr><td>Epistemic</td><td>0.710</td><td>0.419</td><td>0.524</td><td>0.638</td><td>0.442</td><td>0.520</td></tr>
<tr><td>Deontic</td><td>0.629</td><td>0.360</td><td>0.455</td><td>0.573</td><td>0.383</td><td>0.457</td></tr>
<tr><td>Dynamic</td><td>0.388</td><td>0.233</td><td>0.290</td><td>0.391</td><td>0.287</td><td>0.330</td></tr>
<tr><td>Evaluation</td><td>0.592</td><td>0.319</td><td>0.412</td><td>0.523</td><td>0.302</td><td>0.382</td></tr>
<tr><td>Interrogative</td><td>0.844</td><td>0.789</td><td>0.815</td><td>0.818</td><td>0.789</td><td>0.803</td></tr>
<tr><td>Directive</td><td>0.692</td><td>0.309</td><td>0.418</td><td>0.695</td><td>0.354</td><td>0.458</td></tr>
<tr><td>Assertive</td><td>0</td><td>0</td><td>0</td><td>0.1</td><td>0.031</td><td>0.047</td></tr>
<tr><td>Commissive</td><td>0.83</td><td>0.277</td><td>0.409</td><td>0.713</td><td>0.155</td><td>0.246</td></tr>
<tr><td>MacroAvg</td><td>0.540</td><td>0.296</td><td>0.343</td><td>0.522</td><td>0.306</td><td>0.347</td></tr>
<tr><td>Accuracy</td><td>0.620</td><td></td><td></td><td>0.621</td><td></td><td></td></tr>
</table>

Table 9: 5-fold cross validation result of finer-grained classification with f_1 and f_2 features.

for the classification. Besides the parsing problem, there are other linguistic issues behind. Many modal operators could result in different modalities, such as 应该 *ying1gai1* "should", 会 *hui4* "will/can/may", 要 *yao4* "want/will/should/must" etc. Sometimes, it is hard to decide which meaning is correct in a context. There may be also other linguistic issues that we have not discovered yet. This corpus thus could be used for both linguistic study and computational applications, e.g. event processing.

5 Conclusion

In this paper, we present a Chinese corpus annotated with modalities, speech acts and finer-grained event types. We also provide experiments on classification in different levels of categories with a general feature set. The experimental result is acceptable concerning the difficult linguistic issues behind. In future, we would like to continue our research work on improving the corpus and exploring more semantic information including lexical semantic structures and lexical relations such as WordNet to improve the performance of the classification.

Acknowledgements

The work is supported by a General Research Fund (GRF) sponsored by the Research Grants Council (Project no. 543810 and 543512).

References

John Langshaw Austin. 1975. *How to do things with words: Second Edition*. Harvard University Press, Cambridge, MA.

Luciana Beatriz Avila and Heliana Mello. 2013. Challenges in modality annotation in a brazilian portuguese spontaneous speech corpus. *Proceedings of WAMM-IWCS2013*.

Kathryn Baker, Michael Bloodgood, Bonnie Dorr, Chris Callison-Burch, Nathaniel Filardo, Christine Piatko, Lori Levin, and Scott Miller. 2012. Use of modality and negation in semantically-informed syntactic mt. *Language in Society*, 38(2).

Farah Benamara, Baptiste Chardon, Yannick Mathieu, Vladimir Popescu, and Nicholas Asher. 2012. How do negation and modality impact on opinions? In *Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics*, pages 10–18.

Defang Cao, Wenjie Li, Chunfa Yuan, and Kam-Fai Wong. 2006. Automatic chinese aspectual classification using linguistic indicators. *International Journal of Information Technology*, 12(4):99–109.

Chih-Chung Chang and Chih-Jen Lin. 2011. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):1–27.

Pi-Chuan Chang, Huihsin Tseng, Dan Jurafsky, and Christopher D. Manning. 2009. Discriminative reordering with chinese grammatical relations features. In *Proceedings of the Third Workshop on Syntax and Structure in Statistical Translation*, pages 51–59.

Keh-Jiann Chen and Chu-Ren Huang. 1990. Information-based case grammar. In *Proceedings of the 13th conference on Computational linguistics*, pages 54–59.

Keh-Jiann Chen, Chu-Ren Huang, Li-Ping Chang, and Hui-Li Hsu. 1996. Sinica corpus: Design methodology for balanced corpora. In *Proceedings of Pacific Asia Conference on Language, Information and Computing (PACLIC)*, pages 167–176.

Chu-Ren Huang, Feng-Yi Chen, Keh-Jiann Chen, Zhao ming Gao, and Kuang-Yu Chen. 2000. Sinica treebank: design criteria, annotation guidelines, and on-line interface. In *Proceedings of the second workshop on Chinese language processing: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics*, pages 29–37.

Alexis Palmer, Elias Ponvert, Jason Baldridge, and Carlota Smith. 2007. A sequencing model for situation entity classification. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 896–903.

Frank Robert Palmer. 2001. *Mood and Modality*. Cambridge University Press, Cambridge.

Roser Sauri and James Pustejovsky. 2012. Are you sure that this happened? assessing the factuality degree of events in text. *Computational Linguistics*, 38(2):261–299.

Roser Sauri, Marc Verhagen, and James Pustejovsky. 2006. Annotating and recognizing event modality in text. In *Proceedings of 19th International FLAIRS Conference*, pages 333–338.

John R. Searle. 1976. A classification of illocutionary acts. *Language in Society*, 5(1):1–23.

Eric V. Siegel and Kathleen R. McKeown. 2000. Learning methods to combine linguistic indicators: Improving aspectual classification and revealing linguistic insights. *Computational Linguistics*, 26(4):595–628.

Eric V. Siegel. 1999. Corpus-based linguistic indicators for aspectual classification. In *Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics*, pages 112–119.

Carlotta Smith. 1991. *The Parameter of Aspect*. Kluwer Academic Publishers, Dordrecht.

Huihsin Tseng, Pichuan Chang, Galen Andrew, Daniel Jurafsky, and Christopher Manning. 2005. A conditional random field word segmenter for sighan bakeoff 2005. In *Proceedings of the Fourth SIGHAN Workshop on Chinese Language Processing*, volume 171.

Zeno Vendler, 1967. *Linguistics in Philosophy*, chapter Verbs and times, pages 97–121. Cornell University Press, Ithaca.

Hongzhi Xu and Chu-Ren Huang. 2013. Primitives of events and the semantic representation. In *Proceedings of the 6th International Conference on Generative Approaches to the Lexicon*, pages 54–61.

Alessandra Zarcone and Alessandro Lenci. 2008. Computational models for event type classification in context. In *Proceedings of the International Conference on Language Resource and Evaluation (LREC)*, pages 1232–1238.

Xiaodan Zhu, Chunfa Yuan, Kam-Fai Wong, and Wenjie Li. 2000. An algorithm for situation classification of chinese verbs. In *Proceedings of the second workshop on Chinese language processing: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics*, volume 12, pages 140–145.